UNIVERSIDAD
DE GRANADA

Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación
Facultad de Ciencias

Doble Grado en Ingeniería Informática y Matemáticas

trabajo de fin de grado

# Optimización de redes neuronales

Presentado por:
Blanca Cano Camarero

Tutores:
Juan Julián Merelo Guervós
*Arquitectura y tecnología de computadores*

Francisco Javier Meri de la Maza
*Análisis matemático*

Curso académico 2021-2022

# Optimización de redes neuronales

Blanca Cano Camarero

Blanca Cano Camarero *Optimización de redes neuronales*.
Trabajo de fin de Grado. Curso académico 2021-2022.

**Responsables de tutorización**

Juan Julián Merelo Guervós
*Arquitectura y tecnología de computadores*

Francisco Javier Meri de la Maza
*Análisis matemático*

Doble Grado en Ingeniería Informática y Matemáticas

Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación
Facultad de Ciencias

Universidad de Granada

# Resumen

Objetivo inofrmática: Elegir un framework común para trabajar con redes neuronales así como una serie de problemas de complejidad media, tales como spambase. Establecer una línea base examinando los resultados obtenidos con la configuración base a la hora de entrenar este tipo de redes neuronales y el resultado obtenido. A partir de esa línea base, testear las diferentes restricciones, cambios en representación y suposiciones deducidos en la parte matemática para ver qué influencia tienen en la velocidad, en el resultado, o en ambos.

Objetivo matemáticas: El objetivo de esta parte es doble, en primer lugar, se propone analizar con detalle las demostraciones de algunos resultados de aproximación universal de redes neuronales para funciones continuas. En segundo lugar se propone realizar un estudio de la posible optimización de redes neuronales concretas en base a los resultados obtenidos empíricamente en la parte informática. Se tratará de modelizar matemáticamente dichos resultados y de obtener mejoras en la convergencia de las aproximaciones imponiendo, si es necesario, hipótesis más restrictivas en algunos de los elementos de las redes neuronales que se correspondan con su uso en la práctica.

Libros: [1] Abu-Mostafa, Y.S. et al.: Learning From Data. AMLBook, 2012. [2] G. Cybenko, Approximations by superpositions of a sigmoidal function, Math. Contro Signal Systems 2 (1989), 303-314. [3] J. Conway, A Course in Functional Analysis, 2nd Edition, Springer-Verlag, 1990. [4] A. Géron, Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems (2nd ed.). O'Reilly, 2019. [5] K. Hornik, M Stinchcombe and H. White, Multilayer feedforward networks are universal approximators, Neural Networks 2 (1989), 359-366. [6] W. Rudin, Real and complex analysis. McGraw-Hill Book Co., New York-Toronto, Ont.-London 1966.

**PALABRAS CLAVE:** redes neuronales LSTM series temporales selección de modelos validación selección de hiperparámetros detección de anomalías detector perturbación

# Summary

**KEYWORDS:** neural networks LSTM time series model selection validation hyper-parameters selection anomaly detection detector perturbation

# Índice general

# Índice de figuras

# Índice de tablas

# 1 Introducción

[ Polinomios de Bernstein] Dada cierta función $f : [0,1] \to \mathbb{R}$, se define el n-ésimo polinomio de Bernstain para $f$ como

$$B_b(x) = B_n(x;f) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k(1-x)^{n-k}.$$

# 2 Polinomios de Bernstein

[Polinomios de Bernstein] Dada cierta función $f : [0,1] \to \mathbb{R}$, se define el n-ésimo polinomio de Bernstain para $f$ como

$$B_n(x) = B_n(x;f) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\left(\binom{n}{k}\right) x^k(1-x)^{n-k}.$$

La intuición que se esconde tras esta definición es la siguiente: Se pretende aproximar la función $f$ a través de los puntos $\frac{k}{n}$ con $n\mathbb{N}$ fijo y $k \in \{0,...,n\}$. Se prentende que si evaluamos $B_n(x)$ con $x$ próximo a un valor de la forma $\frac{k}{n}$ su valor se aproxima al de $f\left(\frac{k}{n}\right)$.

Para ello recordaremos el teorema del Binomio de Newton:

[Binomio de Newton] Cualquier potencia de un binomio $x+y$ con $x, y \in R$, puede ser expandido en una suma de la forma

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

Así pues en virtud de esta igualdad y puesto que nuestro dominio de definición de $f$ es $[0,1]$, para cualquier $x \in [0,1]$.

Tenemos

$$1 = (x + (1-x))^n = \sum_{k=0}^{n} \binom{n}{k} x^k(1-x)^{n-k}$$

Multiplicamos ahora en ambos lados por $f(x)$

$$f(x) = \sum_{k=0}^{n} f(x)\binom{n}{k} x^k(1-x)^{n-k}$$

Y tenemos que las diferencia entre $f(x)$ y $B_n(x)$ es

$$f(x) - B_n(x) = \sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right)\right)\binom{n}{k} x^k(1-x)^{n-k}$$

Que en valores absolutos resulta

$$|f(x) - B_n(x)| = \sum_{k=0}^{n} \left|f(x) - f\left(\frac{k}{n}\right)\right|\binom{n}{k} x^k(1-x)^{n-k}$$

Observando esta ecuación **??** se desprende como es natural un teorema de convergencia.

[Teorema de aproximación d3e Bernstein]

Sea $f$ una función continua en un intervalo $I$ con imágen en los reales. La secuencia de polinomio de Berstein **??** converge uniformemente a $f$ en $I$.

Por estar $f$ definida en un intervalo cerrado, está acotada y además es uniformemente continua.

## Agradecimientos

Agradezco a